



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

BULLETIN
OF THE
TORREY BOTANICAL CLUB

NOVEMBER, 1914

Further observations on the relationship between the number of
ovules formed and the number of seeds developing in *Cercis*

J. ARTHUR HARRIS

(WITH FOUR TEXT FIGURES)

I. INTRODUCTORY REMARKS

In an earlier paper,* I discussed upon large masses of data drawn from series of trees in different habitats the relationship between the number of seeds maturing and the number of ovules formed in the legume, *Cercis canadensis*. The analysis of the more homogeneous collections from individual trees was then reserved. This is now undertaken.

II. ANALYSIS OF DATA

A. *The Meramec Highlands Collections*

For convenience of treatment merely, I recognize two series, the first comprising 10 trees from which a relatively large number of pods were taken and the second embracing 100 trees from which 50-100 pods each were gathered.†

Table I gives the data and Table II the essential constants for the 13 large samples.

That the samples differ from tree to tree is especially conspicuous in the averages. Means such as 3.56, 3.75, 3.88, 4.26,

* Harris, J. Arthur. On the relationship between the number of ovules formed and the number of seeds developing in *Cercis*. Bull. Torrey Club 41: 243-256. 1914.

† From three trees of this latter collection a much larger number of pods was taken; they are, therefore, included here. The first hundred pods are also treated below, where the small samples from individual trees are discussed.

[The BULLETIN for October (41: 483-532) was issued 28 O 1914.]

TABLE I
OVULES AND SEEDS DEVELOPING PER POD

Number of Individual	1/2	2/2	1/3	2/3	3/3	1/4	2/4	3/4	4/4	1/5	2/5	3/5	4/5	5/5	1/6	2/6
1	18	32	24	72	90	18	49	137	144	2	13	25	35	28	1	—
2	1	2	1	9	5	2	18	31	35	4	14	36	37	45	—	2
3	2	3	3	14	12	3	18	45	49	1	3	12	29	38	—	—
4	6	14	9	45	66	3	28	45	81	—	—	3	5	13	—	—
5	4	15	4	63	117	7	61	156	212	—	7	11	28	12	—	—
6	—	—	—	—	—	—	—	1	1	—	—	6	5	7	—	3
7	—	2	2	12	19	3	31	72	95	1	10	29	63	53	—	2
8	2	11	17	45	95	13	44	140	204	3	3	13	36	40	—	1
9	3	11	6	20	12	10	31	77	48	6	48	137	210	213	8	31
10	—	1	—	4	1	—	6	10	8	3	1	10	30	30	—	1
115	—	1	2	6	12	1	15	41	69	1	12	23	58	85	—	—
116	—	—	—	2	3	—	11	50	40	—	16	81	159	126	—	3
118	—	—	—	2	5	—	3	11	36	—	—	13	50	95	—	—

3/6	4/6	5/6	6/6	2/7	3/7	4/7	5/7	6/7	7/7	2/8	4/8	5/8	6/8	7/8	8/8	Totals
—	2	5	—	—	—	—	—	—	—	—	—	—	—	—	—	695
4	12	10	5	—	—	—	—	—	1	—	—	—	—	—	—	274
—	—	5	3	—	—	—	—	1	—	—	—	—	—	—	—	242
—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	318
—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	698
10	6	18	10	—	1	5	8	11	4	1	—	4	—	—	—	111
4	9	5	6	—	—	—	—	—	—	—	—	—	—	—	—	418
1	1	1	2	—	—	—	—	—	—	—	—	—	—	—	—	672
97	219	328	209	6	16	27	59	61	40	—	1	3	1	3	1	1,942
2	5	13	18	1	—	—	3	2	3	—	—	—	—	—	—	152
6	17	20	30	—	—	—	—	1	—	—	—	—	—	—	—	400
45	78	136	82	1	3	4	10	20	9	—	—	—	—	1	—	880
1	23	70	104	—	1	1	6	16	14	—	—	—	—	—	—	451

TABLE II
PHYSICAL CONSTANTS FOR THIRTEEN INDIVIDUALS

Plant	Pods	Ovules			Seeds		
		Average	Standard Deviation	V	Average	Standard Deviation	V
1	695	3.760 ± .021	.832 ± .015	22.12	2.935 ± .026	1.023 ± .018	34.87
2	274	4.671 ± .033	.811 ± .023	17.37	3.555 ± .046	1.142 ± .033	32.14
3	242	4.268 ± .035	.812 ± .025	19.02	3.496 ± .048	1.114 ± .034	31.88
4	318	3.563 ± .027	.710 ± .019	19.92	2.965 ± .036	.966 ± .026	32.57
5	698	3.768 ± .016	.638 ± .011	16.94	3.129 ± .021	.835 ± .015	26.68
6	111	6.153 ± .052	.808 ± .036	13.12	4.585 ± .075	1.174 ± .053	25.60
7	418	4.409 ± .024	.743 ± .017	16.86	3.555 ± .032	.985 ± .023	27.70
8	672	3.887 ± .018	.689 ± .013	17.74	3.231 ± .025	.969 ± .018	30.00
9	1,942	5.542 ± .016	.931 ± .011	16.79	4.281 ± .022	1.276 ± .016	29.80
10	152	5.131 ± .050	.908 ± .035	17.70	4.230 ± .070	1.285 ± .050	30.37
115	400	4.765 ± .028	.821 ± .019	17.24	4.012 ± .037	1.099 ± .026	27.38
116	880	5.375 ± .018	.782 ± .012	14.55	4.293 ± .024	1.070 ± .017	24.93
118	451	5.466 ± .027	.856 ± .019	15.65	4.913 ± .032	1.001 ± .022	20.37

TABLE III

DIFFERENCES AND PROBABLE ERRORS OF DIFFERENCES IN THE STANDARD DEVIATION
OF OVULES AND SEEDS AND IN THE CORRELATION OF OVULES AND SEEDS
FOR SELECTED INDIVIDUALS OF CERCIS

Individuals Compared	Ovules		Seeds		Ovules and Seeds	
	Differences in Standard Deviation	Ratio	Differences in Standard Deviation	Ratio	Differences in Correlation	Ratio
1 and 5	.193 $\pm$.019	10.28	.189 $\pm$.024	7.93	.071 $\pm$.025	2.77
7 and 9	.187 $\pm$.021	9.04	.291 $\pm$.028	10.47	.053 $\pm$.028	1.93
8 and 10	.219 $\pm$.037	5.87	.316 $\pm$.053	6.00	.089 $\pm$.036	2.44
4 and 118	.146 $\pm$.027	5.42	.035 $\pm$.034	1.02	.073 $\pm$.027	2.72
5 and 9	.292 $\pm$.016	18.04	.441 $\pm$.022	20.32	.048 $\pm$.025	1.91
6 and 10	.100 $\pm$.051	1.98	.111 $\pm$.073	1.52	.302 $\pm$.065	4.66

TABLE IV

RELATIONSHIP BETWEEN NUMBER OF OVULES PER POD AND NUMBER OF SEEDS PER
POD

Tree	Number of Pods	C. F.	r_{os}	r_{oz}	$r_{oz}/E_{r_{oz}}$
1	695	.7807	.573 $\pm$.017	-.074 $\pm$.025	2.93
2	274	.6709	.488 $\pm$.031	-.060 $\pm$.058	1.03
3	242	.8189	.679 $\pm$.023	+.112 $\pm$.063	1.77
4	318	.8323	.648 $\pm$.022	+.047 $\pm$.056	.82
5	698	.8304	.502 $\pm$.019	-.152 $\pm$.025	6.08
6	111	.7452	.342 $\pm$.056	-.178 $\pm$.092	1.94
7	418	.8062	.497 $\pm$.025	-.128 $\pm$.032	3.94
8	672	.8311	.556 $\pm$.018	-.043 $\pm$.026	1.64
9	1,942	.7724	.550 $\pm$.012	-.016 $\pm$.015	1.03
10	152	.8243	.645 $\pm$.032	+.081 $\pm$.080	1.01
115	400	.8420	.607 $\pm$.021	-.028 $\pm$.034	.82
116	880	.7987	.531 $\pm$.016	-.062 $\pm$.023	2.73
118	451	.8989	.720 $\pm$.015	-.070 $\pm$.032	2.19

4.76, 5.13, 5.54 and 6.15 with probable errors always in the second place of decimals and ranging from .016 to .057 for ovules are so clearly significantly different that it is needless to calculate probable errors. The differences for mean number of seeds are also clearly significant.

The standard deviations for both ovules and seeds present a problem more difficult of solution by mere inspection. Values like .63, .68, .71, .74, .78, .85 and .93 with probable errors not exceeding .04 in any case appear to be significant. Differences exceeding their probable error as widely as those for the few random pairs given in Table III certainly indicate that the trees must be regarded as individual in variability as well as type.

For both ovules and seeds the coefficients of variation differ

conspicuously, ranging from 13.12 to 22.12 for ovules and from 20.37 to 34.87 for seeds. Unless the values of means and standard deviations are perfectly correlated—an assumption which we have no reason to make—one would expect the coefficients of variation to show considerable fluctuation in magnitude from the influence of the means alone.

Since the constants showing the mean values and the variabilities of ovules and seeds per pod and perhaps those for the correlation of these two characters as well (see Table III) differ significantly from individual to individual it is clear that there may be serious disadvantages in lumping together the materials from different trees to form a large sample to be used in the investigation of delicate correlations.

Turn now to the correlation coefficients. Table IV gives the results. Here as in the two random samples and their combination there is a considerable correlation between the number of ovules per pod and the actual number of seeds developing.

Trees 1, 5, 7, 8, and 9 have been selected for the test of linearity of regression because of the large number of pods available. The equations to the regression straight lines are given in the

Tree	Equation to Regression Straight Line	Deviation*
1	$s = .284 + .705\ o$	.065
5	$s = .653 + .657\ o$	.042
7	$s = .652 + .658\ o$	.050
8	$s = .194 + .781\ o$	.030
9	$s = .100 + .754\ o$	.048

accompanying table. The average weighted deviation of observation from theory of each tree is also given. Graphs have been made showing the empirical means and the fitted straight line for each tree. Only that for tree 9 need be given, as figure 1. Perhaps the fit is slightly better than for some of the others, but the observations are numerous. All of the graphs show unusually good agreements of theoretical and empirical means. The theoretical and the empirical means differ on an average by only about five-hundredths of a seed.

* Weighted mean deviation, disregarding signs, of observed from theoretical mean number of seeds per pod.

The coefficients for the correlation between the number of ovules per pod and the capacity of the pods for developing their ovules into matured seeds as determined from the formula

$$r_{os} = \frac{r_{os} - v_o/v_s}{\sqrt{1 - r_{os}^2 + (r_{os} - v_o/v_s)^2}}$$

are all very low and are either positive or negative in sign, for these individual trees — presumably more homogeneous than the

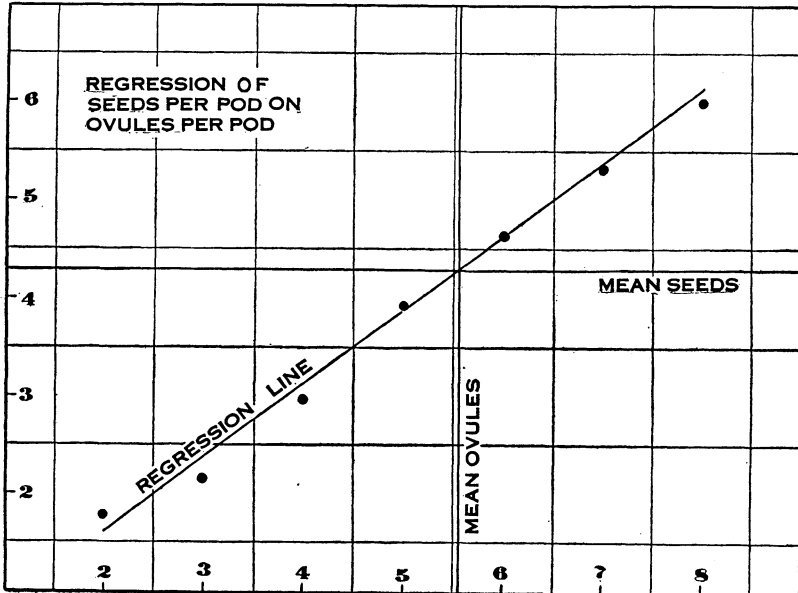


FIG. 1. Regression line and empirical means for tree 9.

general samples. Regarding probable errors, only four of the thirteen can be considered to differ significantly from 0. Ten of the thirteen are negative. All those which are significant with regard to their probable error are negative in sign.

Smaller collections from 100 trees were gathered shortly after those described above for the purpose of having more numerous individual samples for testing the results already stated. Separate correlation tables showing the relationship between the number of ovules formed and the number of seeds developing were drawn up for each tree. Of course it is impossible to publish all these.

The constants which interest us here are (a) the coefficient of

correlation for the number of ovules formed and the number of seeds developing per pod, and (b) the coefficient of correlation between the number of ovules per pod and the deviation of the number of seeds per pod from the probable number on the assumption that the number of seeds per pod is proportional to the number of ovules. These constants are presented in Table V in a form which will be clear without further comment.

The values for all the trees (large and small collections) are seriated in Table VI.

At the outset of the work I had thought that perhaps by the collection of a sufficiently wide series of individuals I might find some in which the coefficient of correlation between the number of ovules formed and the number of seeds developing would be very low. The results for the 110 trees show, however, that not only are all the values positive but that in every case they are of a substantial order. The lowest value entered in Table VI is .325 and the highest .850.

The constants of these correlation coefficients calculated with Sheppard's correction, are:

Mean,	.5994 $\pm$.0075
Standard Deviation,	.1165 $\pm$.0053
Coefficient of Variation,	19.431

It seems of considerable interest for our present problem to try to ascertain whether this observed variation in the magnitudes of the correlations between the number of ovules and seeds represents a real biological difference in the individual trees examined, or whether it is merely a statistical result, attributable to the fact that constants for each tree are based upon a small, not immensely large, sample of its pods. I believe we may make some progress as follows.

The standard deviation of random sampling of r is $(1 - r^2)/\sqrt{n}$, where n is the number of individuals included in the sample on which r is calculated. Let us assume now that the true value of r is the same for each individual tree and that it has the value found as our mean, say $r = .600$. This seems reasonable.

If n for each individual were 100, a condition holding for 60 of the trees but only with rough approximation for the remaining 50, one would expect

TABLE V
CORRELATIONS FOR SERIES OF PODS FROM INDIVIDUAL TREES

Tree	N	r_{os}	r_{oz}	Tree	N		r_{oz}	Tree	N	r_{os}	r
11	100	.336	-.245	45	100	.562	.077	79	100	.663	-.073
12	100	.342	-.291	46	100	.630	.064	80	100	.655	-.070
13	100	.444	-.275	47	87	.716	.283	81	100	.506	-.253
14	100	.455	-.098	48	97	.442	-.248	82	73	.707	.019
15	100	.609	.151	49	93	.463	-.156	83	83	.597	-.201
16	100	.578	-.014	50	100	.576	-.077	84	76	.598	-.095
17	100	.704	.121	51	86	.728	.197	85	100	.683	-.010
18	100	.665	.157	52	100	.510	-.120	86	100	.720	.122
19	100	.477	-.164	53	100	.720	.152	87	69	.569	.004
20	100	.527	-.090	54	68	.522	-.022	88	100	.702	.154
21	88	.618	.065	55	100	.596	-.009	89	100	.647	.072
22	85	.460	-.051	57	100	.522	-.278	90	91	.709	.116
23	100	.681	.172	58	100	.688	-.062	91	100	.633	-.032
24	100	.684	.111	59	86	.752	.282	92	100	.486	-.036
25	100	.510	-.186	60	100	.641	-.116	94	80	.591	-.011
26	100	.691	.127	61	100	.732	-.029	95	97	.717	.081
27	100	.662	.048	62	98	.747	.112	96	100	.680	-.025
28	100	.618	-.106	63	74	.661	.011	98	74	.551	.045
29	100	.593	-.048	64	79	.630	.030	100	65	.490	-.185
30	91	.604	-.031	65	100	.808	-.070	101	100	.545	-.107
31	80	.436	-.034	66	80	.644	-.080	102	99	.560	-.171
32	85	.659	.053	67	70	.698	-.088	107	100	.440	-.192
33	95	.429	-.220	68	100	.716	.050	108	100	.600	-.010
34	100	.357	-.189	69	73	.789	-.106	109	69	.500	-.026
35	83	.437	-.053	70	100	.461	-.132	111	67	.428	-.198
36	100	.560	-.058	71	100	.678	.086	112	80	.579	.004
37	90	.563	.055	72	100	.803	.012	113	72	.742	.119
38	100	.645	.083	73	100	.750	.016	114	100	.475	-.004
39	100	.745	.167	74	100	.662	-.047	115	100	.807	.150
40	88	.667	.041	75	80	.659	-.063	116	100	.506	.090
41	76	.723	.226	76	100	.747	-.085	117	100	.393	-.166
42	100	.343	.011	77	100	.596	-.078	118	100	.688	-.106
43	62	.454	.012	78	94	.707	-.189	119	50	.854	.085
44	100	.618	.097								

TABLE VI
DISTRIBUTION OF COEFFICIENTS OF CORRELATION, r_{os}

Center of Correlation Grade	f	Center of Correlation Grade	f
.325	1	.600	9
.350	4	.625	6
.375	—	.650	12
.400	1	.675	9
.425	4	.700	9
.450	7	.725	8
.475	4	.750	6
.500	9	.775	—
.525	3	.800	4
.550	6.5	.825	—
.575	6.5	.850	1

$$\sigma_r = \frac{1 - .60^2}{\sqrt{100}} = .0640.$$

This would have its probable error

$$.67449 \times .0640/\sqrt{220} = .0029.*$$

By comparison we get:

$$\text{Empirical, S. D.} = .1165 \pm .0053,$$

$$\text{Theoretical, S. D.} = .0640 \pm .0029,$$

$$\text{Difference} = .0515 \pm .0060.$$

The difference is over 8 times its probable error and I think indicates that there are real biological differences in the individuals.

The constants for the correlation between the number of ovules per pod and the deviation of the number of seeds from their probable value, r_{oz} , for the 60 trees with 100 pods per tree and for the 40 other trees with less than 100 pods are summarized from Table V in Table VII. For a grand total of 110 trees I add

TABLE VII
FREQUENCIES OF VALUES OF r_{oz} FOR SERIES OF INDIVIDUALS

Ovules and Capacity, r_{oz}	60 Trees, 100 per Tree	40 Trees, Less Than 100	Total 110 Trees
-.325 - -.275	2	—	2
-.275 - -.225	3	1	4
-.225 - -.175	3	5	9
-.175 - -.125	3	2	7
-.125 - -.075	10	4	14
-.075 - -.025	11	6	20
-.025 - +.025	8	7	16
+ .025 - +.075	4	6	11
+ .075 - +.125	8	5	15
+ .125 - +.175	8	—	8
+ .175 - +.225	—	1	1
+ .225 - +.275	—	1	1
+ .275 - +.325	—	2	2

the constants obtained for the 10 individuals from which all the pods were taken.

These results are also shown graphically in Fig. 2. Here the length of the lines indicates the magnitude of the correlation, i. e., the amount by which it deviates from 0, and the nature and

* The 10 trees yielding more than 100 pods each have been included.

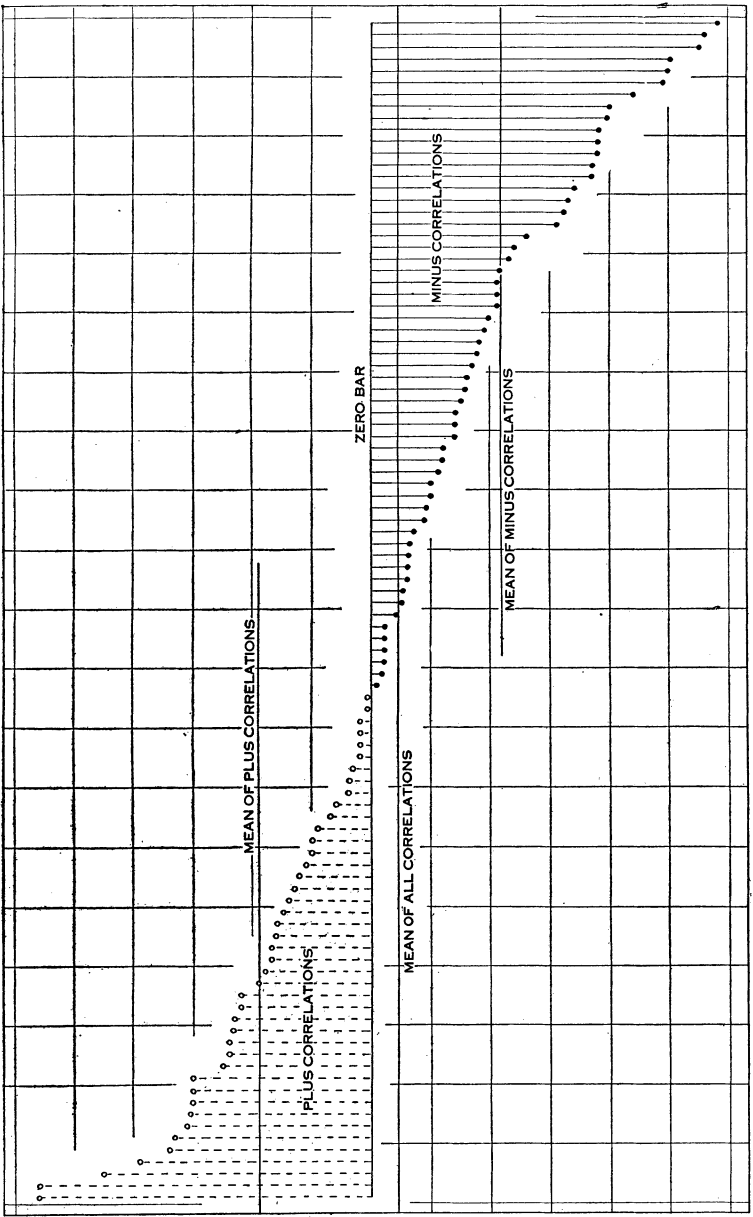


FIG. 2. Values of r_s for 100 individual trees from Meramec Highlands. Each coefficient is represented by a line. For r_s space is equal to .050.

the direction of this line, the sign. The firm lines and solid dots below the zero bar stand for negative coefficients, while the broken lines and circles above the zero bar indicate the magnitude of positive coefficients. The ten large trees are omitted.

Taking these results as they stand it would appear that the relationship between the number of ovules per pod and the capacity of the pod for maturing its seeds may vary, being sometimes positive and sometimes negative. The light and dark line areas in the diagram are *nearly* equal, and their means lie *about* equally distant from the zero bar.

But one cannot accept any statistical constant—average standard deviation or coefficient of correlation—as absolutely correct as a description of the material from which the sample investigated was drawn; all are too large or too small by an amount known to mathematicians as the probable error of random sampling. Is it not possible that the considerable range of variation in the constants tabled here is due to purely statistical causes and has no biological significance whatever? I think we may proceed as follows.

Assuming that there is no relationship between the number of ovules per pod and the capacity of the pod for maturing its seeds—i. e., that $r_{oz} = 0$ —the deviation of the empirical means and the empirical standard deviation due to the probable errors of random sampling from the theoretical 0 may be determined. I illustrate with the 60 trees furnishing each 100 pods.

The standard deviation of the coefficient of correlation is as pointed out above $(1 - r^2)/\sqrt{n}$. Clearly where $n = 100$ and r_{oz} is actually 0 one would expect a standard deviation of 0.10 due merely to the errors of random sampling. But this standard deviation itself would have a probable error which for the 60 trees from which I have 100 pods would be $.67449(0.1/\sqrt{120})$, or .0062.

One must expect, therefore, if r_{oz} be actually 0, to find a standard deviation of $0.1000 \pm .0062$ in the coefficients for the 60 trees with 100 pods each *due to no organic cause whatever but solely to the errors of random sampling*.

The probable error of the mean is $.67449\sigma_r/\sqrt{n}$. Substituting values with $\sigma = 0.1$ as just indicated I find

$$E_a = .67449 \times 0.1/\sqrt{60} = .0087.$$

This line of argument has been followed out in the preparation of Table VIII where for convenience I assumed that $n = 100$ for each tree.* In this table, the constants actually found are compared with the values one should expect them to have if the corre-

TABLE VIII
MEAN VALUES OF r_{02} FOR THE INDIVIDUAL TREES

	60 Trees	40 Trees	110 Trees
Theoretical mean.....	.0000 $\pm$.0087	.0000 $\pm$.0107	.0000 $\pm$.0064
Calculated mean.....	-.0292 $\pm$.0106	-.0175 $\pm$.0154	-.0232 $\pm$.0080
Difference.....	-.0292 $\pm$.0137	-.0175 $\pm$.0187	-.0232 $\pm$.0102
Theoretical standard deviation	.1000 $\pm$.0062	.1000 $\pm$.0075	.1000 $\pm$.0045
Calculated standard deviation.	.1218 $\pm$.0075	.1440 $\pm$.0108	.1242 $\pm$.0056
Difference.....	.0218 $\pm$.0097	.0440 $\pm$.0131	.0242 $\pm$.0072

lation between the number of ovules per pod and their capacity for maturing their seed were actually 0, and the constants as found from actual collections were due merely to the errors of random sampling.

For all three series the mean differs from 0 by less than 2.5 times its probable error. For the standard deviations the difference between the observed and the theoretical values is less than 2.5 in the case of the 60 trees and only about 3.3 times its probable error in the other two series.

Notwithstanding these low values of the means and the nearly equal areas of plus and minus values on the diagram, one must not lose sight of the facts, (a) that there are more negative than positive coefficients, the ratio being 57 : 43 for the smaller collections and 64 : 46 for the series from all the individual trees, and (b) that all the means are negative in sign.

This series like those discussed above seems to indicate that if there is any relationship between the number of ovules per ovary and the capacity of the ovary for maturing its ovules into seeds it is of such a nature that the ovaries with the larger numbers of ovules are slightly less capable than those with the smaller numbers.

* This is not strictly true; I think that the approximation is quite close enough for present purposes. Sheppard's correction was used for the empirical distribution.

B. The Individuals from the Vicinity of Lawrence, Kansas

The relationship for the whole material has already been discussed. For the individual trees the results as shown in Table IX are most interesting. Of the 22 values of r_{oz} , only 1 has

TABLE IX
CORRELATIONS FOR INDIVIDUALS

Tree	r_{os} and $E r_{os}$	r_{oz} and $E r_{oz}$
1	.671 $\pm$.037	— .134 $\pm$.066
2	.692 $\pm$.035	+ .099 $\pm$.099
3	.595 $\pm$.044	— .052 $\pm$.067
4	.604 $\pm$.043	— .204 $\pm$.065
5	.580 $\pm$.045	— .172 $\pm$.065
6	.549 $\pm$.047	— .131 $\pm$.066
7	.400 $\pm$.057	— .182 $\pm$.065
8	.380 $\pm$.058	— .189 $\pm$.065
9	.365 $\pm$.058	— .341 $\pm$.060
10	.365 $\pm$.058	— .319 $\pm$.061
11	.554 $\pm$.047	— .174 $\pm$.065
12	.394 $\pm$.057	— .189 $\pm$.065
13	.609 $\pm$.042	— .030 $\pm$.067
14	.391 $\pm$.057	— .263 $\pm$.063
15	.381 $\pm$.058	— .274 $\pm$.062
16	.543 $\pm$.048	— .117 $\pm$.067
17	.537 $\pm$.048	— .253 $\pm$.063
18	.181 $\pm$.065	— .322 $\pm$.060
19	.572 $\pm$.045	— .022 $\pm$.067
20	.644 $\pm$.039	— .212 $\pm$.064
21	.514 $\pm$.050	— .085 $\pm$.067
22	.635 $\pm$.040	— .288 $\pm$.062

the positive sign. Taking the ratios of the constants to their probable errors to test their significance I find that the single positive constant deviates from 0 by an amount equal to its probable error; hence no significance can be attached to it. Of the 21 negative coefficients 14 differ from 0 by more than 2.5 times their probable error. The whole situation is summed up graphically in diagram 3 in which these ratios are plotted out with the signs of the correlations. If there were no real biological correlation between the number of ovules per pod and their capacity for maturing their seeds, the distribution of these ratios would be centered at 0, marked by the "theoretical mode and mean" line, with the deviations about equally distributed above

and below. What one does find is that the empirical mode falls on the class -3.50 to -2.50 . The area of the polygon which is composed of statistically significant constants is shaded in.

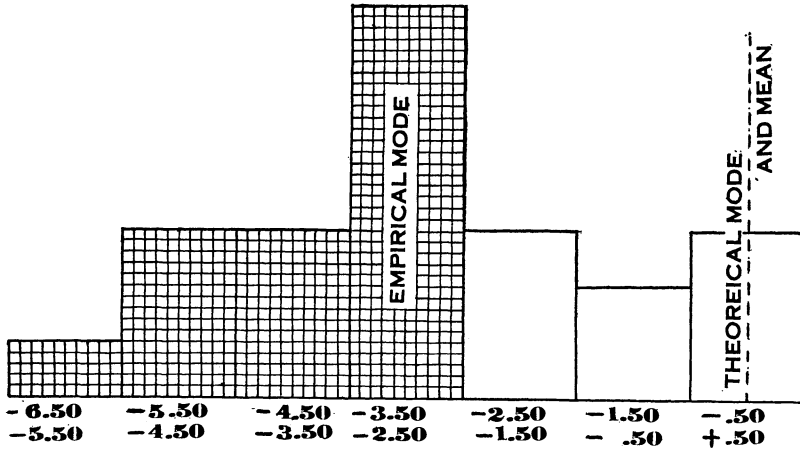


FIG. 3. Distribution of ratios of r_{oz} coefficients to their probable errors, showing large number of significantly negative correlations for the eastern Kansas series.

C. *The Individuals from the Vicinity of Sharpsburg, Ohio*

In this collection, the lumped data gave values of $r_{os} = .455 \pm .009$, $r_{oz} = -.034 \pm .011$. Thus the correlation between number of ovules and the capacity of the ovary for maturing its ovules into seeds while negative in sign is not only low but is only 3.15 times its probable error.

The calculated correlations for the individuals are set forth in Table X.

These results differ essentially from those secured for the Kansas series in two regards; (a) only 6 out of the 26 may be considered statistically significant with regard to their probable error as compared with 14 out of 21 in the Kansas series; (b) the constants are about evenly distributed between positive and negative, there being 14 positive and 12 negative signs. The general average is, however, negative.

Again, the ratios of the constants to their probable errors are plotted out in a polygon (FIG. 4) showing the scatter of the constants on either side of zero.

TABLE X
CORRELATIONS FOR INDIVIDUALS

Tree	r_{os} and $E_{r_{os}}$	r_{oz} and $E_{r_{oz}}$
I	.511 ± .041	− .079 ± .055
2	.790 ± .021	+ .078 ± .055
3	.566 ± .037	+ .022 ± .055
4	.614 ± .034	+ .205 ± .055
5	.310 ± .050	− .066 ± .055
6	.581 ± .036	+ .184 ± .055
7	.396 ± .046	+ .083 ± .055
8	.426 ± .045	− .057 ± .055
9	.254 ± .051	+ .020 ± .055
10	.425 ± .045	− .058 ± .055
11	.324 ± .049	− .060 ± .055
12	.485 ± .042	+ .056 ± .055
13	.614 ± .034	− .074 ± .055
14	.119 ± .054	− .164 ± .054
15	.524 ± .040	+ .047 ± .055
16	.567 ± .037	+ .108 ± .054
17	.400 ± .046	− .280 ± .051
18	.454 ± .044	+ .071 ± .055
19	.440 ± .044	+ .041 ± .055
20	.342 ± .049	+ .137 ± .054
21	.338 ± .049	− .183 ± .053
22	.470 ± .043	− .062 ± .055
23	.614 ± .034	− .075 ± .055
24	.416 ± .045	− .064 ± .055
25	.386 ± .047	− .089 ± .055
26	.262 ± .051	− .086 ± .055

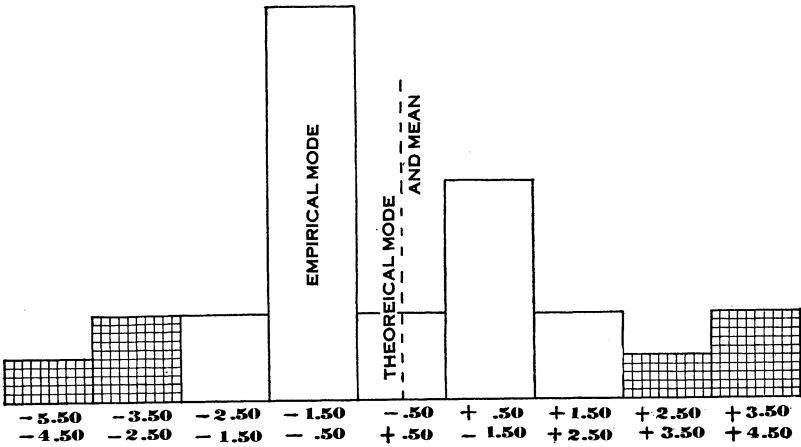


FIG. 4. Distribution of ratios of r_{oz} to $E_{r_{oz}}$, Ohio series.

III. SUMMARY AND DISCUSSION

The investigations described in this and the preceding paper establish several points concerning fertility and fecundity. The following may now be cited.

(a) The physical constants — type, variability, and correlation—of the number of ovules per pod and the number of seeds developing per pod in *Cercis canadensis* differ sensibly from individual to individual and from habitat to habitat. The data do not, however, justify the conclusion that the trees from the different habitats are to be distinguished taxonomically.

(b) The correlations for number of ovules formed and number of seeds developing per pod, r_{os} , have always been found positive and of a moderate, considerable or even high intensity.

This is true for the pods of an individual tree as well as for a mixed sample from a considerable series of trees. The correlation coefficient is slightly raised by the combination of collections from different individuals.

(c) Regression is sensibly linear, both within the series of pods from the same individual and in a population of pods from many individuals. Possibly, however, there is a departure from linearity in the pods with eight ovules, but in my largest series there are only 36 of these pods out of a total of 28,554; this number is too small to be given great importance.

The significance of the linearity of regression is two-fold. Statistically, it justifies describing the interdependence between the number of ovules formed and the number of seeds maturing by the coefficient of correlation. Biologically, it shows that the rate of increase in number of seeds developing per pod remains the same as we pass from pods with the lowest to pods with the highest numbers of ovules.

(d) Wherever large series of pods have been examined, the correlation between the number of ovules per pod and the capacity of the pods for maturing their seeds, r_{oz} , has a negative sign and a low, usually a very low, magnitude. When the number of pods is relatively small — say about 100 as in the case of the correlations from individual trees — the coefficient is sometimes positive. These results may well be due to the probable errors of random sampling which, with samples of this small size, may be quite large enough to screen such a slight relationship.

In such cases the number of negative values is generally larger than the number of positive coefficients, and their mean numerical magnitude is always higher. For every large series examined the value of r_{oz} has been over 2.5 times its probable error and sometimes many times its probable error. These evidences can leave little doubt of the existence of a slight negative relationship between the number of ovules formed and the capacity of the pod for maturing its ovules into seeds, the pods with the larger number of ovules producing relatively fewer seeds.

This conclusion has also been reached in an earlier paper for the dwarf varieties of *Phaseolus vulgaris* as a whole.

(e) The foregoing conclusions and other statements made in these pages apply exclusively to the one species considered and should not be extended to others except on the basis of actual data. There is no reason to assume that species may not differ in this regard. The data available for another of the Leguminosæ, *Robinia*,* indicate that quite different conditions from those found in *Cercis* may prevail. If the correlations found for *Sanguinaria*† are based on sufficiently large and representative samples they lead to the same conclusion. There are strong evidences that some strains of *Phaseolus* differ from others in the sign of this relationship. Indeed the Kansas series of *Cercis* differs rather conspicuously from others in the intensity of the negative correlation.

The conclusions concerning capacity for seed development here drawn are based upon mature pods only. One of the most important things to be done is to determine the relation of this phenomenon to the intra-individual selective elimination of ovaries, if it occurs in *Cercis*. All of the data here discussed were collected before this differential failure of ovaries in *Staphylea* was demonstrated. As yet I have been unable to obtain adequate materials for solving the problem for *Cercis*.

(f) This paper is exclusively a statement of observed facts. I have no explanation to offer of the relationships which have been regularly found when adequately large series of data have been analyzed. Theories as to the causes underlying the conditions

* Harris, J. Arthur. *Biometrika* 6: 441-442. 1909.

† Harris, J. Arthur. *Biometrika* 7: 321-324. 1910.

observed seem to me, in view of the numerous difficulties of the problem, premature. Upon the painstaking collections of facts in regard to natural phenomena, whether or not they can be lined up with current theories, seems to me to rest the real advance of biology. When more comprehensive data are available—many of which are already collected and in an advanced stage of reduction—it will be much safer to consider causal phases of the phenomena.

COLD SPRING HARBOR